

## THE TRANSVERSE PECULIAR VELOCITY OF THE Q2237+0305 LENS GALAXY AND THE MEAN MASS OF ITS STARS

SHAWN POINDEXTER<sup>1</sup>, CHRISTOPHER S. KOCHANÉK<sup>1</sup>*Draft version February 9, 2010*

## ABSTRACT

Using 11-years of OGLE V-band photometry of Q2237+0305, we measure the transverse velocity of the lens galaxy and the mean mass of its stars. We can do so because, for the first time, we fully include the random motions of the stars in the lens galaxy in the analysis of the light curves. In doing so, we are also able to correctly account for the Earth's parallax motion and the rotation of the lens galaxy, further reducing systematic errors. We measure a lower limit on the transverse speed of the lens galaxy,  $v_t > 338 \text{ km s}^{-1}$  (68% confidence) and find a preferred direction to the East. The mean stellar mass estimate including a well-defined velocity prior is  $0.12 \leq \langle M/M_\odot \rangle \leq 1.94$  at 68% confidence, with a median of  $0.52 M_\odot$ . We also show for the first time that analyzing subsets of a microlensing light curve, in this case the first and second halves of the OGLE V-band light curve, give mutually consistent physical results.

*Subject headings:* gravitational lensing — methods: numerical — quasars: general — quasars: individual (Q2237+0305) — galaxies: kinematics and dynamics

## 1. INTRODUCTION

Quasar microlensing provides a unique tool for studying the properties of cosmologically distant lens galaxies and the structure of quasars (see Wambsganss 2006). Each of the multiple images of the quasar passes through the gravitational potential of the stars along the line-of-sight in the lens galaxy. These stars microlens each of the “macro” images, so the total magnification of each quasar image is strongly affected by the lensing effects of the stars and the size of the quasar emission region. Since the observer, lens galaxy, stars, and source quasar are all moving, these magnifications change on timescales of 1-10 years with order unity amplitudes.

The relevant physical scale for quasar microlensing is the Einstein radius projected into the source plane plane,

$$R_E = D_{\text{OS}} \sqrt{\frac{4G \langle M \rangle}{c^2} \frac{D_{\text{LS}}}{D_{\text{OL}} D_{\text{OS}}}} \\ = 1.8 \times 10^{17} \left( \frac{\langle M \rangle}{M_\odot} \right)^{1/2} \text{ cm}, \quad (1)$$

where  $G$  is the gravitational constant,  $c$  is the speed of light,  $\langle M \rangle$  is the mean stellar mass of the stars,  $D_{\text{LS}}$ ,  $D_{\text{OL}}$ , and  $D_{\text{OS}}$  are the angular diameter distances between the lens-source, observer-lens, and observer-source respectively, where we have used the lens and source redshifts for Q2237+0305 ( $z_l = 0.0394$ ,  $z_s = 1.685$ , Huchra et al. 1985, Q2237 hereafter). The observed microlensing amplitude is controlled by the ratio between the source size,  $R_V \approx 6 \times 10^{15} \text{ cm}$  (in V-band, see our companion paper Poindexter & Kochanek (2009), hereafter Paper II) and  $R_E$ , in the sense that smaller accretion disks produce larger variability amplitudes than larger disks. If a source is much larger than

$R_E$ , there is little change in the magnification.

The timescale for variability is determined by the relative velocities of the observer, the lens, its stars, and the source. Generally, the lens motions dominate (Kayser & Refsdal 1989), leading to two characteristic timescales. There is the timescale to cross an Einstein radius,

$$t_E \approx \frac{R_E}{v_{\text{lens}}} \frac{(1+z_l)D_{\text{OL}}}{D_{\text{OS}}} \\ \approx 8 \left( \frac{v_{\text{lens}}}{462 \text{ km s}^{-1}} \right)^{-1} \left( \frac{R_E}{2 \times 10^{17} \text{ cm}} \right) \text{ yr}, \quad (2)$$

and there is the timescale to cross the source,

$$t_s \approx \frac{R_V}{v_{\text{lens}}} \frac{(1+z_l)D_{\text{OL}}}{D_{\text{OS}}} \\ \approx 0.4 \left( \frac{v_{\text{lens}}}{462 \text{ km s}^{-1}} \right)^{-1} \left( \frac{R_V}{6 \times 10^{15} \text{ cm}} \right) \text{ yr}, \quad (3)$$

where  $v_{\text{lens}}$  is the expected transverse speed of the lens for Q2237. These timescales are also affected by the direction of motion relative to the shear (Wambsganss et al. 1990). Variation is guaranteed on timescale  $t_E$  and can be observed on timescale  $t_s$  if the magnification pattern locally has structure on the scale of  $R_V$ .

Quantitative studies using quasar microlensing have exploded in the last few years. Recent efforts have studied the relationships between accretion disk size and black hole mass (Morgan et al. 2010), size and wavelength (Anguita et al. 2008; Bate et al. 2008; Eigenbrod et al. 2008b; Poindexter et al. 2008; Floyd et al. 2009; Mosquera et al. 2009), sizes of non-thermal (X-ray) and thermal emission regions (Pooley et al. 2007; Morgan et al. 2008; Chartas et al. 2009; Dai et al. 2009), and the dark matter fraction of the lens (Dai et al. 2009; Pooley et al. 2009;

<sup>1</sup> Department of Astronomy and Center for Cosmology and AstroParticle Physics, The Ohio State University, 140 W 18th Ave, Columbus, OH 43210, USA, (sdp,ckochanek)@astronomy.ohio-state.edu

Mediavilla et al. 2009). All these studies have used static magnification patterns which ignore the random stellar motions in the lens galaxy. However, the stellar velocity dispersions of lens galaxies are comparable to the peculiar velocities of galaxies, and ignoring them can lead to biased results. For example, the average direction of motion of the source is the same for all images, but this coherence is limited by the random motions of the stars. With fixed patterns one must either overestimate the coherence, by using the same direction of motion for each image, or underestimate it by using independent directions for each image (see Kochanek et al. 2007). In either case, it would be dangerous to attempt measurements that depend on this coherence, such as disk shape and orientation, or the transverse peculiar velocity of the lens.

Kundic & Wambsganss (1993), Schramm et al. (1993), and Wambsganss & Kundic (1995) considered the effects of random stellar motions and found that these motions can also lead to shorter microlensing time scales because the pattern velocities of the microlensing caustics can be much higher than any physical velocities. As a result, measurements based on static magnification patterns may underestimate source sizes or mean masses or overestimate the transverse velocity in order to match the effects created by the random stellar motions. Wyithe et al. (2000a) showed that it is possible to statistically correct for these effects and that the velocity correction can be up to 40% depending on the optical depth and shear. Another benefit to dynamic magnification patterns is the ability to properly account for the velocity of Earth around the Sun and the rotation of the lens galaxy (Tuntsov et al. 2004).

Dynamic magnification patterns are also important because they impart a well-measured physical scale to the patterns. All direct microlensing observations are in “Einstein units” where the length scale is  $\langle M \rangle^{1/2}$  cm. Determining masses, velocities, or source sizes requires some sort of dimensional prior. In our studies we have generally followed Kochanek (2004) and used velocity priors designed to mimic the combined effects of random and ordered motion. The reason for focusing on the velocity is that we know two of the contributions, our velocity and the stellar velocities, and the remaining peculiar velocities of the lens and source are truly random variables for which we have reasonable priors from cosmological models. Source sizes turn out to be little affected by the choice of priors (see the discussion in Kochanek 2004), but estimates of the true velocity and the mean stellar masses are affected. Hopefully by including the true random stellar motions we can further reduce the sensitivity of microlensing results to such priors.

In this paper we use microlensing to measure the peculiar velocity of a lens galaxy and the mean mass of its stars including the effects of the stellar motions, Earth’s motion, and the rotation of the lens galaxy. The transverse velocity direction can be measured with microlensing because the shear sets a preferred direction for each image and the statistics of variability depend on the motion relative to this axis (see Figure 1). In theory, accurately measuring the transverse peculiar velocity of many galaxies over a broad range of redshifts could form the ba-

sis of a new cosmological test (Gould 1995). Measuring the mean stellar mass, including remnants, is an independent means of checking local accountings (e.g. Gould 2000), which must be assembled from very disparate selection methods for high mass, low mass, evolved and dead, remnant stars. Moreover, doing this is possible in detail only for the Galaxy. While microlensing is relatively insensitive to the mass function (see Paczynski 1986; Wyithe et al. 2000b), there are some prospects of exploring this in the future as well (e.g. Wyithe & Turner 2001; Schechter et al. 2004; Congdon et al. 2007).

This work expands on the methods described in Kochanek (2004) and Kochanek et al. (2007) by adding the random stellar motions in the lens galaxy. In this paper we address the computational issues and then apply this improved technique to determine the transverse motion of Q2337 and the mean mass of its stars. In Paper II, we study the shape of the accretion disk of the source quasar. We describe the photometric data in §2. Then we describe the Bayesian Monte Carlo Method and the models we use in §3. Our results are presented in §4 followed by a discussion in §5. We use an  $\Omega_M = 0.3, \Omega_\Lambda = 0.7$  flat cosmological model with  $H_0 = 72 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

## 2. DATA

The quadruply lensed  $z_s = 1.695$  quasar Q2237 was discovered by Huchra et al. (1985). The images are observed through the bulge of a barred Sab lens galaxy at a projected distance  $\sim 0''.9$  ( $\sim 700 \text{ pc}$ ). The very low  $z_l = 0.0394$  lens redshift leads to very fast lens motions projected onto the source plane, leading to variability timescales as short as  $\sim 0.2$  years (Equations 2 and 3). Microlensing of Q2237 was first observed by Irwin et al. (1989) and confirmed by Corrigan et al. (1991). There are also detailed mass models and dynamical studies by Schneider et al. (1988), Kent & Falco (1988), Rix et al. (1992), Mihov (2001), Trott & Webster (2002), and van de Ven et al. (2008).

We analyze nearly 11 years of the Optical Gravitational Lensing Experiment V-band photometric monitoring data for Q2237 (Udalski et al. 2006). To speed our analysis and as a cross check on the results, we divided the OGLE data into two separate light curves. The first light curve is from JD 2,450,663 to JD 2,452,621 and consists of 100 epochs and will be referred to as LC1. The second light curve has 230 epochs from JD 2,452,763 to JD 2,454,602 and will be referred to as LC2. Each light curve covers just over 5 years. The light curves are shown in Figure 2.

Since Q2237 is expected to have a very small time delay between its images (e.g. Wambsganss & Paczynski 1994), we only need to subtract the light curves (in magnitudes) to remove the intrinsic variability of the quasar. We estimated the systematic photometric errors in the OGLE data using each successive triplet of epochs spanning less than 15 days. We used the first and last point of each triplet to predict the middle observation and then derived the systematic error that, when added in quadrature to the OGLE uncertainties, make the predictions consistent with the uncertainties. These systematic error estimates are 0.02, 0.03, 0.04, and 0.05 magnitudes for images A, B, C, and D respectively.

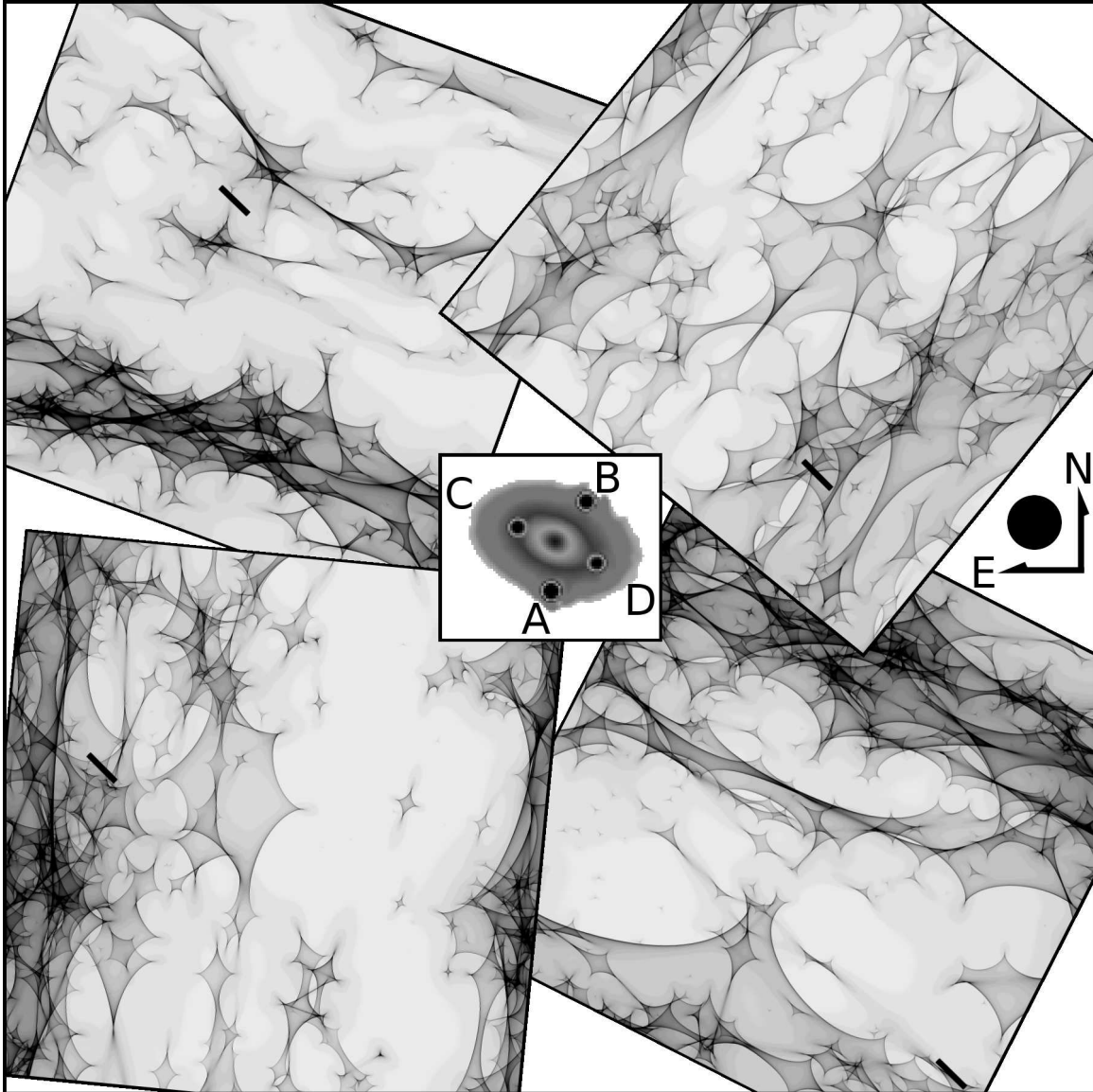


FIG. 1.— Example of a trial source trajectory (dark line segments) superposed on an instantaneous point-source magnification pattern for  $\langle M/M_\odot \rangle = 0.3$ . Darker shades indicate higher magnification. An HST H-band image in the center labels the images and the corresponding magnification patterns. Each pattern is rotated to have the correct orientation relative to the lens. This particular LC2 trial has an effective lens-plane velocity of  $\sim 600 \text{ km s}^{-1}$  Northeast. The solid disk at right has a radius of  $10^{17} \text{ cm}$  on these patterns.

### 3. METHODS

Our Bayesian Monte Carlo method (Kochanek 2004) requires the construction of magnification patterns and a model of the quasar accretion disk. The patterns are convolved with the source model and used to produce large numbers of simulated light curves for comparison to the data. The results for these trial light curves are combined in a Bayesian analysis to measure parameters and their uncertainties. Here we describe the generation of the patterns (§3.1), the source model (§3.2), the Bayesian Monte Carlo method and its priors (§3.3), and the computational techniques needed to allow for stellar motion (§3.4).

#### 3.1. Magnification Patterns

We generate dynamic magnification patterns (see Figure 1 for examples of instantaneous patterns) in the source plane by randomly placing stars near each macro

TABLE 1  
LENS GALAXY MODEL PARAMETERS.

Image	$\kappa$	$\gamma$	PA [deg]
A	0.40	0.40	175
B	0.38	0.39	-39
C	0.73	0.72	70
D	0.62	0.62	-63

NOTE. — The normalized surface density ( $\kappa$ ), shear ( $\gamma$ ) and its position angle at the location at each image.

image in the lens galaxy. The normalized surface density and shear are determined by fitting models to the *HST* astrometry of the four images relative to the lens galaxy and the mid-IR image flux ratios (Agol et al. 2000). We modeled the lens as a power law mass profile with the `lensmodel` program of the `gravlens` package (Keeton 2001). Because all the images are  $\sim 700 \text{ pc}$  in projected

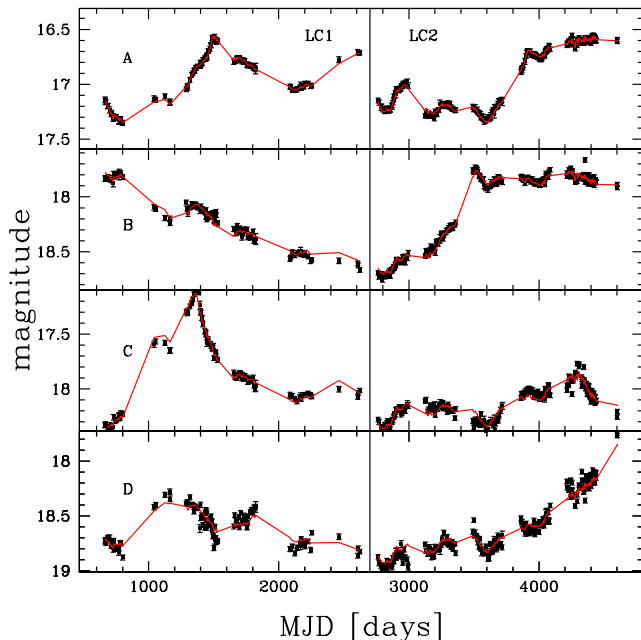


FIG. 2.— The OGLE Q2237 V-band light curves. The left panel is LC1 and the right is LC2. The rows from top to bottom are images A, B, C, and D. Here we show the corrected error bars. The red curves are one of our best fit models for the microlensing and intrinsic source variation. Because we only determine the light curve at the epochs with data, gaps are filled by linear interpolation.

distance from the galactic center, we expect the surface density to be dominated by the stars rather than by dark matter ( $\kappa_*/\kappa = 1$ ). This assumption is corroborated by the microlensing analysis of Kochanek (2004) and the dynamical models of van de Ven et al. (2008). The normalized surface density,  $\kappa$ , and tidal shear,  $\gamma$ , from this model (see Table 1) were used to generate the magnification patterns. The lens plane is populated using a mass function of  $dN/dM \propto M^{-1.3}$  with a dynamic range of  $M_{\max}/M_{\min} = 50$  based on the Galactic mass function of Gould (2000). We use patterns with mean masses of  $\langle M \rangle = 0.01, 0.03, 0.1, 0.3, 1, 3.0$ , and  $10 M_{\odot}$ . We use the Kochanek (2004) particle-particle/particle-mesh implementation of the inverse ray-shooting method (e.g. Kayser et al. 1986) to create  $N_{\text{pix}}^2 = 4096^2$  magnification patterns. The shear is slightly adjusted (by at most  $1/N_{\text{pix}}$ ) in order to produce periodic magnification patterns that eliminate edge effects, as detailed in the Appendix of Kochanek (2004). The outer scale of the patterns is  $20 R_E = 3.7 \times 10^{18} \langle M/M_{\odot} \rangle^{1/2}$  cm, which results in a resolution of  $0.005 R_E/\text{pixel} = 9.0 \times 10^{14} \langle M/M_{\odot} \rangle^{1/2}$  cm/pixel. For comparison, Morgan et al. (2010) estimate that the black hole mass corresponds to a gravitational radius of  $r_g = GM_{\text{BH}}/c^2 = 2 \times 10^{14}$  cm and in Paper II we find a disk scale length of  $6 \times 10^{15}$  cm (half-light radius of  $1.5 \times 10^{16}$  cm).

For the first time in any model of microlensing data, we fully include the random motions of the microlenses by using an animated sequence of magnification patterns. We use the measured one-dimensional velocity dispersion of  $170 \text{ km s}^{-1}$  (van de Ven 2009, personal communication, also Trott et al. (2008); Foltz et al. (1992)). We assign

each star a random velocity as a Gaussian random deviate of amplitude  $\sigma_* = 170 \text{ km s}^{-1}$  for each coordinate and then generate a magnification pattern for each image/epoch combination. While binary stars make up a large fraction of stellar systems (e.g. Fischer & Marcy 1992), they should not have a significant effect on the patterns. Only relatively close binaries ( $\ll 1 \text{ AU}$ ) have significant orbital velocities compared to the stellar or bulk motions, and such separations are very small compared to the Einstein radius ( $1100 \text{ AU}$  for  $1 M_{\odot}$  in the lens plane). Thus, binary motion is only significant for close binaries, but close binaries have separations much smaller than the Einstein radius or our estimated source sizes ( $66 \text{ AU}$  for  $1 M_{\odot}$  in the lens plane, see Paper II) and would be indistinguishable from a single point mass on our patterns. In effect, binaries should only act like a shift in the mean mass of the stars.

### 3.2. Disk Model

We employ a generic thin disk model for which the surface temperature scales as  $T_s \propto R^{-3/4}$  with radius  $R$  (Shakura & Sunyaev 1973). The microlensing signal is primarily sensitive to the half-light radius of the disk (Mortonson, Schechter & Wambsganss 2005), so the details of the radial profile have limited effects. As in Kochanek (2004), we neglect the inner disk edge, since it should have few observed effects given disk sizes at these wavelengths. We define the area of the disk to be the area enclosed within the contour defined by  $kT = hc/\lambda$ . We first parametrize the source models by choosing from 24 different projected areas covering  $\log_{10}(\text{area}/\text{cm}^2) = 29.2$  to  $33.8$  in steps of  $0.2$ . For each source area, we used five inclinations,  $i$ , with a  $\cos i$  of  $0.2, 0.4, 0.6, 0.8$ , and  $1.0$  (face-on), and for each area and inclination, we used 18 different equally spaced position angles for the major axis of the disk. Paper II discusses the disk model in detail.

### 3.3. The Bayesian Monte Carlo Method

We use the Bayesian Monte Carlo Method of Kochanek (2004). We randomly generate light curves from the animated microlensing magnification patterns over the full range of physical parameters and source sizes and fit them to the observed light curves. We then use Bayes Theorem to infer the likelihood distribution of the parameters given the fit statistics for the light curves. Each simulated light curve is defined by

$$m_i(t) = S(t) + \mu_i + \delta\mu_i(t) + \Delta\mu_i = S(t) + \mu_{i,\text{tot}}, \quad (4)$$

where  $S(t)$  is the intrinsic variability of the source,  $\mu_i$  is the macro model magnification,  $\delta\mu_i(t)$  is the microlensing, and  $\Delta\mu_i$  is the systematic magnification offset for each image,  $i$ . For each trial we compute the goodness of fit

$$\chi^2 = \sum_i \sum_t \left( \frac{m_i(t) - S(t) - \mu_{i,\text{tot}}}{\sigma_i(t)} \right)^2 \quad (5)$$

after solving for the optimal model of the source variability  $S(t)$  and the magnification offsets. The parameters we vary in this study include the projected area of the disk, the inclination of the disk, the position angle of the disk, the effective velocity of the source, and the mean mass of the stars in the lens galaxy. We call these the physical parameters,  $\xi_p$ . For any combination of these

parameters, we also randomly select starting points on each of the magnification patterns and refer to these nuisance parameters,  $\xi_n$ .

We calculate the likelihood of the data given the parameters as

$$P(LC|\xi_p, \xi_n) \propto \Gamma \left[ \frac{N_{\text{dof}}}{2}, \frac{\chi^2}{2} \right], \quad (6)$$

where  $\Gamma$  is the incomplete Gamma function. Kochanek (2004) justifies this form by allowing for uncertainties in the magnitude errors,  $\sigma_i(t)$  and averaging over these uncertainties. This ensures that the likelihood is consistent for trials fitting better than  $\chi^2/\text{dof} = 1$  given the formal uncertainties.

The probability of the parameters given the data is then

$$P(\xi_p, \xi_n|LC) \propto P(LC|\xi_p, \xi_n)P(\xi_p)P(\xi_n), \quad (7)$$

where  $P(\xi_p)$  and  $P(\xi_n) = 1$  are the prior probability distributions of the physical and nuisance parameters. Since we are analyzing two separate parts of the same light curves, we combine the results to improve our measurements by multiplying the probabilities for each light curve and then applying the priors,

$$P(\xi_p|LC1, LC2) \propto P(LC1|\xi_p)P(LC2|\xi_p)P(\xi_p). \quad (8)$$

We did this for two reasons. First, it becomes (probably exponentially) harder to find good fits to longer and longer light curves. Second, analyzing the curves separately allows us to study whether different light curves for the same object give the same answers. The price is that analyzing them separately and then combining them will have less statistical power than a simultaneous analysis of all the data. We compute the probability distributions by marginalizing over the nuisance variables

$$P(\xi_p|LC) \propto \int P(\xi_p, \xi_n|LC)d\xi_n. \quad (9)$$

We compute this as a Monte Carlo integration over the trial light curves, which should converge to the true integral if we generate enough simulated light curves.

For each source size, inclination, and disk position angle we must first convolve the magnification pattern with the source model. Then we produce trial light curves by choosing random starting points and velocities across the animated sequence of magnification patterns. In addition to the random motions of the stars, we must also assign bulk velocities to the observer,  $\mathbf{v}_o$ , lens,  $\mathbf{v}_l$ , and source,  $\mathbf{v}_s$ , leading to an effective (source-plane) velocity of

$$\mathbf{v}_e = \frac{\mathbf{v}_o}{1+z_l} \frac{D_{LS}}{D_{OL}} - \frac{\mathbf{v}_l}{1+z_l} \frac{D_{OS}}{D_{OL}} + \frac{\mathbf{v}_s}{1+z_s} \quad (10)$$

(e.g. Kayser et al. 1986) that is dominated by the lens velocity,  $\mathbf{v}_l$ , in the case of Q2237. From the projection of the CMB dipole (Hinshaw et al. 2009), we know that  $\mathbf{v}_o = (-50, -23) \text{ km s}^{-1}$  East and North respectively for Q2337. Based on Tinker, Wetzel, & Zehavi (2009) we estimate that the (1D) rms peculiar velocities of the lens and source are  $\sigma_{\text{lens}} = 327 \text{ km s}^{-1}$  and  $\sigma_{\text{src}} = 230 \text{ km s}^{-1}$ , respectively. For the calculations, we randomly draw each effective velocity coordinate (in the

lens plane) from a one dimensional Gaussian distribution with  $\sigma = 1000 \text{ km s}^{-1}$  and then re-weight the trials to a more physical range when we carry out the Bayesian integrals.

As discussed in the introduction, quasar microlensing is subject to a degeneracy between mean stellar mass, effective velocity, and accretion disk size. Including random stellar motions partially breaks this degeneracy by introducing a physical scale to the magnification patterns. We still need a prior on one of these parameters to make useful measurements. As in Kochanek (2004), we apply a velocity prior. Here we define our prior in the lens-plane, since the lens motion dominates the effective velocity. In the absence of any “streaming velocities”, we can determine the peculiar velocities only up to a  $180^\circ$  degeneracy that corresponds to a time reversal symmetry given that the peculiar velocity priors depend on the speed but not the direction of motion. Our velocity prior in the lens-plane is

$$P(\mathbf{v}_l) \propto \exp \left( -\frac{(\mathbf{v}_l - \mathbf{v}_{\text{CMB},l})^2}{2\sigma^2} \right), \quad (11)$$

where  $\mathbf{v}_{\text{CMB},l}$  is our CMB motion projected onto the lens-plane, and the expected dispersion in the lens-plane is

$$\begin{aligned} \sigma^2 &= \sigma_{\text{lens}}^2 + (\sigma_{\text{src}}(1+z_l)D_{OL}/D_{OS})^2 \\ &= (327 \text{ km s}^{-1})^2. \end{aligned} \quad (12)$$

The very high projected motion of the lens due to the very low lens redshift means that the source motion is unimportant even though  $\sigma_{\text{src}} \sim \sigma_{\text{lens}}$ . “Streaming velocities”, such as our motion relative to the CMB, our orbit around the Earth (parallax effect), or rotational velocities in the lens break this degeneracy. These motions (up to  $\sim 10\%$  of the peculiar velocities), do slightly break this degeneracy.

With the stars moving it is also makes sense to include the effects of the Earth’s motion and the small rotation velocity of the lens galaxy as part of the motions across the animated patterns. Aside from Tuntsov et al. (2004) the motion of the Earth has not previously been included in a quasar microlensing calculation. Earth’s motion projected onto the lens plane is approximately 10% of the expected transverse velocity of the lens motion (the dominant motion of the system). It is also  $\sim 20\%$  of the minimum possible velocity scale set by the random stellar motions. In trials with  $\langle M/M_\odot \rangle = 0.3$ , we found that including parallax increased the total probability of all trials by  $\sim 20\%$ , and reversing the Earth’s motion reduced the probability by a similar amount. Earth’s orbit is trivial to include and computationally inexpensive, so we include it in our standard calculations even though it has modest effects on the likelihood.

The lens is an late type galaxy with rotation in the plane of the sky of  $\sim 55 \text{ km s}^{-1}$  for images A and B, and  $\sim 20 \text{ km s}^{-1}$  for images C and D (Trott et al. 2008). The position angle of these rotation velocities are  $84.5^\circ$ ,  $-129^\circ$ ,  $-20^\circ$ , and  $153^\circ$  (north through east) respectively for images A, B, C, and D. These are relatively low compared to the disk because the images lie in the bulge. The velocities for images A and B are greater than the modulations introduced by the Earth’s orbit, so we include them in the simulation.

### 3.4. Computational Techniques

The Monte Carlo method requires simultaneous random access to every magnification pattern for each image and epoch. With the stars moving, this means we need 400 and 920 patterns for LC1 and LC2 respectively, corresponding to 25 and 57.5 gigabytes of storage for  $4096^2$  patterns instead of the 4 patterns and 0.25 gigabytes needed for stationary stars. This is more memory than is generally available on any one machine.

Our first step towards solving this computational problem is to conserve memory by compressing the “gray” scale of the convolved patterns. Normally we store patterns as a  $4096^2$  array of 4-byte floating point numbers. However, magnification patterns typically span a dynamic range of 10 magnitudes even for the smallest source sizes, and the data uncertainties are no smaller than 0.01 magnitudes. Thus we only need a logarithmic dynamic range of  $10/0.01 = 1000$  rather than the  $2^{32}$  dynamic range of a floating point variable. For example, if we use 16 bits for each pixel, which is more than sufficient given the dynamic range of the data, we can pack 4 pixels into one 64 bit word rather than using 128 bits, which not only compresses the data by a factor of 2 but also has advantages for data transfer speeds. In practice, we adjust the compression level for each magnification pattern and source size to have a resolution at least ten times better than the uncertainty in the corresponding data point. We achieve compression ratios of 2.5 to 3 for the OGLE data.

Even after compression, the full collection of magnification patterns is still too large for most single machines, so we distribute them evenly among parallel computers. This has the added benefit of utilizing additional CPUs, but at the cost of needing to communicate between nodes. Our goal is to minimize the need for this communication. We sort each light curve in chronological order and then distribute the epochs in a round robin fashion to each node, so that each node has a sparse but complete representation of the data. Trial light curves are started on the individual nodes. If a trial’s  $\chi^2$  exceeds a threshold it is simply discarded. If it is under the threshold, it is passed to other nodes to be tested against the rest of the light curve. This basically amounts to a low resolution pre-search for good-fitting light curves before doing any expensive communication with the other nodes. These light curves are optimized by exploring slightly different starting points and velocities across the magnification patterns. This requires the master node for each trial to do many communications with the other nodes to compute the full  $\chi^2$ , but our tests show that this finds good fits faster than trying more light curves.

For each source model, we choose  $10^5$  starting points and velocities for one of the four images. Then we search for pairwise matches by trying  $10^4$  starting points on each of the other three images and keep the best match for each image. A light curve is then produced from this velocity and starting point for each image. The  $\chi^2$  for each trial light curve is computed from the data. If the  $\chi^2$  of a trial exceeds a threshold during its calculation, we discard it immediately. Such poor solutions will make no contribution to the Bayesian integrals (Equation 9), so there is no point in wasting further calculation or communication on completing the trial.

LC1 was processed in 1.6 CPU-years and LC2 was processed in 2.8 CPU-years utilizing 16 AMD Opteron machines (64 processor cores) simultaneously at the Ohio Supercomputing Center. In total we tried  $9 \times 10^{14}$  unique starting points and  $3 \times 10^9$  different velocities. We found significantly fewer good fitting trials for LC1, so our threshold for saving trials was  $\chi^2/\text{dof} < 4$  for LC1 and  $\chi^2/\text{dof} < 2.5$  for LC2. Our best fit simulated light curves have  $\chi^2/\text{dof} = 0.86$  for LC1, and  $\chi^2/\text{dof} = 0.99$  for LC2. There was a large event in image C, and more rapid magnification changes in LC1 as compared to LC2 (see Figure 2), and this likely explains why it was harder to find good fits for LC1. With these cuts,  $3 \times 10^6$  and  $6 \times 10^6$  trials fits passed the cuts for LC1 and LC2 and were saved. Even though the best fitting light curves produced  $\chi^2/\text{dof} \sim 1$ , we rescaled the  $\chi^2$  for each light curve to produce better-defined results in the Monte Carlo integral (Equation 9) for each parameter. We divided the  $\chi^2$  of trials of LC1 and LC2 by 1.9 and 1.4 respectively for our final analysis, so that  $10^4$  trials for each set were less than  $\chi^2/\text{dof}$  after rescaling. In general, this is conservative and broadens the parameter uncertainties by  $\sqrt{1.9}$  and  $\sqrt{1.4}$  over what we should achieve with an infinite number of trials.

## 4. RESULTS

Here we estimate the transverse peculiar velocity of the lens galaxy, the mean mass of its stars, and the mean magnification offsets defined in Equation 4. We quote the results from the combined analysis of LC1 and LC2, but also show the results from the independent analyses of LC1 and LC2 in Figures 3, 4, 5, and 6. Our main results are based on the  $\sigma = 327 \text{ km s}^{-1}$  lens-plane (1D) peculiar velocity prior described in §3.3 and assume that all the light comes directly from the accretion disk. We verify the latter assumption and examine the structure and orientation of the accretion disk in Paper II.

### 4.1. Transverse velocity

Figure 3 shows the likelihood contours for the transverse velocity of the lens galaxy. The peak likelihood is  $\sim 300 \text{ km s}^{-1}$  East. The individual light curve results are very consistent with the joint analysis, as shown by superposing the 68% contours for LC1 and LC2 in Figure 3. This agreement can also be seen in the position angle of the lens motion (Figure 4) and in the lens speed (Figure 5).

After integrating over direction we find a transverse speed of  $438^{+253}_{-213} \text{ km s}^{-1}$  ( $438^{+419}_{-325} \text{ km s}^{-1}$ ) at 68% (95%) confidence (Figure 5) including our standard prior (Equations 11 and 12). The inclusion of a physical model of the stellar motions does not completely eliminate degeneracies, and our speed estimate is dominated by our velocity prior (Equation 11) at large speeds. If we instead use the broader lens-plane prior with  $\sigma = 1000 \text{ km s}^{-1}$  from which we derive our trials, the measured speed becomes  $1048^{+640}_{-486} \text{ km s}^{-1}$  at 68% confidence basically following the prior. Therefore, we only have measured a lower limit  $v_t > 338 \text{ km s}^{-1}$ . Fixing the mean microlensing mass has little effect, provided the mass is sufficiently large. If we fix the mean mass to be  $\langle M \rangle = 0.3 M_\odot$  we find that  $v_t < 486$  (757)  $\text{km s}^{-1}$  at 68% (95%) confidence.

These results are consistent with earlier results. Wyithe et al. (1999) found a 95% upper limit on the



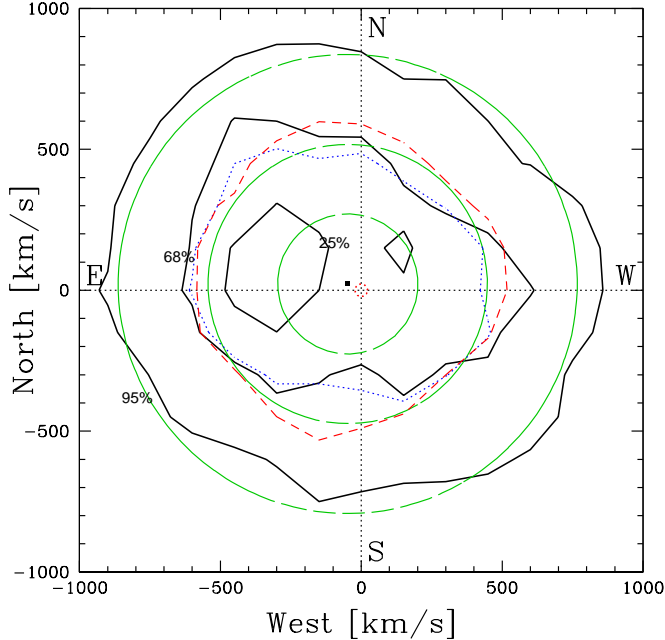


FIG. 3.— Probability distribution for the effective lens-plane velocity. The dark solid contours enclose 25%, 68%, and 95% of the likelihood relative to the peak. Because the motion is strongly dominated by the lens, this is essentially the transverse peculiar velocity of the lens galaxy. The green dashed contours are the velocity prior (Equation 11), drawn at the same levels. The small, red, dotted circle is the 68% contour for the contribution to the prior from the expected motion of the source. The black point is our CMB motion projected onto the lens-plane. The 68% enclosed probability for the LC1 and LC2 analyses are shown as dotted blue and dashed red contours.

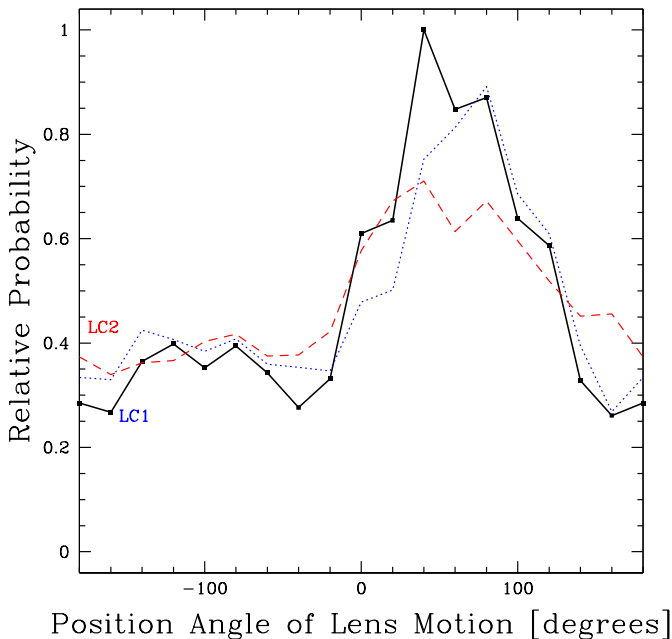


FIG. 4.— Position angle of the effective velocity in the lens-plane. The dotted blue (dashed red) curve shows the results from only the LC1 (LC2) analysis.

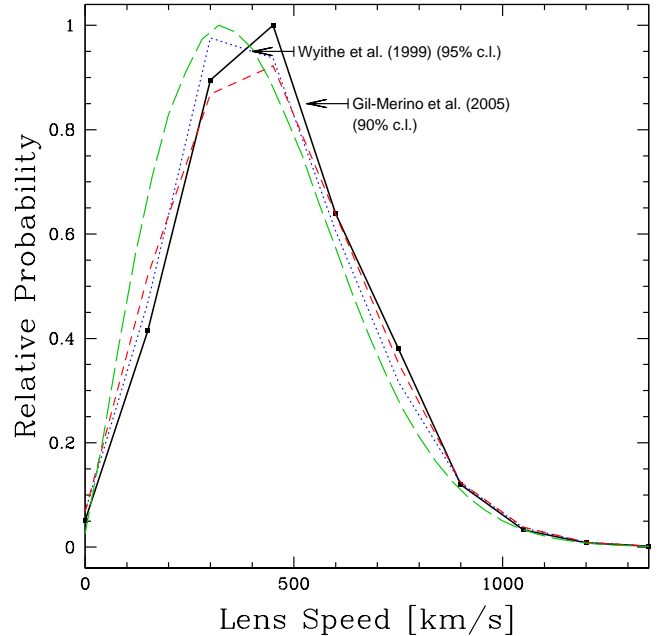


FIG. 5.— Speed distribution of the effective velocity of the source in the lens-plane (dark solid curve). This motion is dominated by the priors on the transverse peculiar velocity of the lens. The green dashed curve shows the prior (Equation 11) integrated over direction. The upper limits from other studies are indicated with arrows (Wyithe et al. 1999; Gil-Merino et al. 2005). The dotted blue (dashed red) curve shows the results from only the LC1 (LC2) analysis.

transverse lens speed,  $v_t < 500 \text{ km s}^{-1}$  by using derivatives of the microlensing light curve from patterns produced from a Salpeter mass function including stellar motions and a range of source sizes. They tried three different mass models with  $\langle M/M_\odot \rangle = 0.22, 0.31$  and  $1.0$  and found that the estimated transverse speed scaled with  $\sqrt{\langle M \rangle}$ . Using the distribution of gaps between high magnification events, but not including stellar motions, Gil-Merino et al. (2005) found a 90% upper limit,  $v_t < 630(2160) \text{ km s}^{-1}$  for  $M = 0.1(1.0)M_\odot$  lenses. It must be noted that neither study included the full range of physical uncertainties we include here.

#### 4.2. Mass

We measure the mean stellar mass in the lens to be  $0.12 < \langle M/M_\odot \rangle < 1.94$  ( $0.04 < \langle M/M_\odot \rangle < 3.46$ ) at 68% (95%) confidence with a median of  $\langle M/M_\odot \rangle = 0.52$ . This is generally consistent with earlier estimates for this lens based on less data and simpler analyses including fewer of the physical uncertainties. It is marginally consistent with the earlier estimate by Kochanek (2004) of  $\langle M/M_\odot \rangle = 0.018^{+0.080}_{-0.015}$  ( $0.018^{+0.270}_{-0.017}$ ) at a 68% (95%) confidence using a similar velocity prior but with less data and static stars. Our uncertainties are a factor  $\sim 2$  times smaller. Lewis & Irwin (1996) argue for a mean mass of  $0.1 < M/M_\odot < 10$  by comparing the observed magnification probability distribution to that of simulations that did not include random stellar motions. Wyithe et al. (2000b) estimate that  $\langle M/M_\odot \rangle = 0.29$  with a lower limit of  $\langle M/M_\odot \rangle \gtrsim 0.11$  at 99% confidence by analyzing the distribution of light curve derivatives and assuming a 1D stellar dispersion of

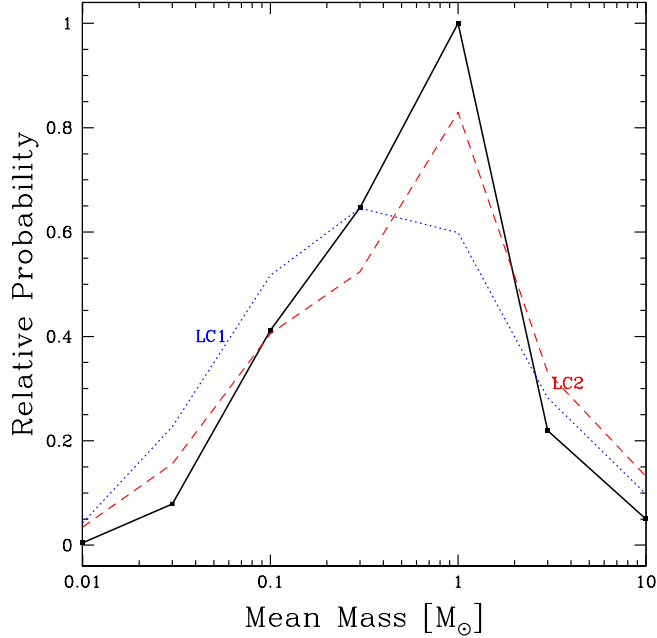


FIG. 6.— Mean mass of the stars. The dotted blue (dashed red) curve shows the results from only the LC1 (LC2) analysis.

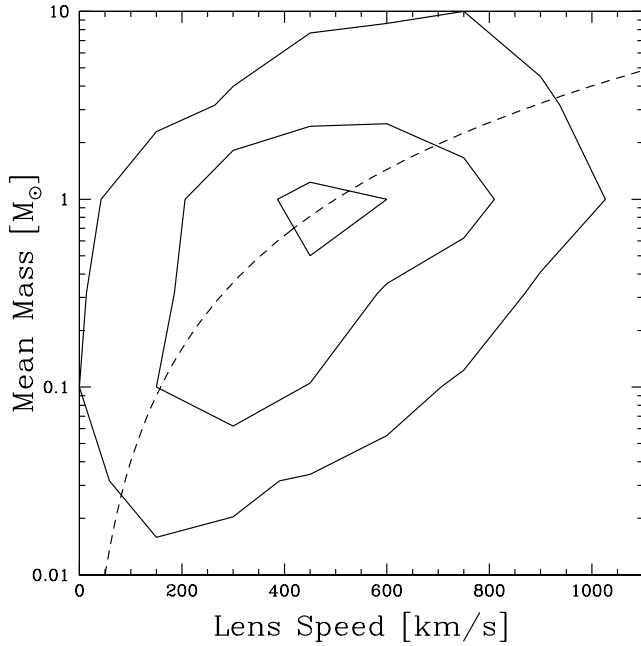


FIG. 7.— Mean mass,  $\langle M \rangle$  of the stars versus the speed of the lens,  $v$ . The dashed line is  $M \propto v^2$ .

$165 \text{ km s}^{-1}$  combined with a prior on the transverse velocity. Gil-Merino & Lewis (2005) simply argued that the masses must be greater than Jupiter-like objects, contrary to the claims of Lee et al. (2005). In our results, the mass estimate is still strongly affected by our velocity prior (Equation 11). If we use the broad  $\sigma = 1000 \text{ km s}^{-1}$  lens-plane velocity prior, the median rises to  $\langle M/M_\odot \rangle = 1.5$ , and we would need to expand the calculations to higher masses to fully sample the mass

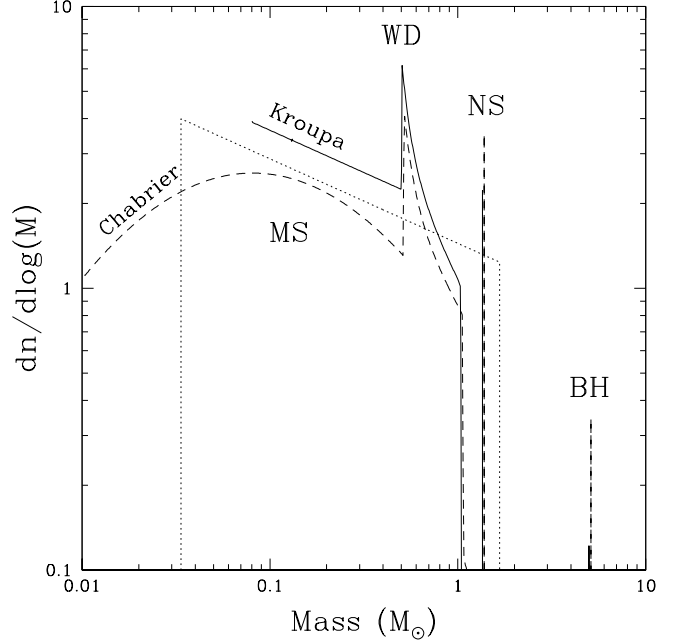


FIG. 8.— The most probable  $\langle M \rangle = 0.3 M_\odot$  microlensing mass function used in the calculations (dotted curve) as compared to that predicted from the Chabrier (2003) or Kroupa et al. (1993) initial mass functions truncated at masses lower than  $0.01 M_\odot$  (mean mass  $0.20 M_\odot$ ) and  $0.08 M_\odot$  (mean mass  $0.32 M_\odot$ ), respectively, for an age of 10 Gyr. The features due to the main sequence (MS), white dwarfs (WD), neutron stars (NS) and black holes (BH) are marked. The Chabrier (2003) and Kroupa et al. (1993) mass functions are slightly offset to make the different amplitudes for neutron stars and black holes visible. The assignment of a  $5 M_\odot$  mass for black holes is arbitrary but not important.

distribution. The problem for accurate mass measurements is that  $\langle M \rangle \propto v^2$  (see Figure 7), so the mean mass is very sensitive to the speed distribution (see Figure 7). Including the mean stellar motions eliminates low velocity solutions corresponding to low masses, but cannot eliminate high mass, high velocity solutions.

We can compare this mass estimate to that expected from stellar mass functions. We approximate the lifetimes of stars by the time to reach the base of the giant branch, using the approximate expression in Hurley et al. (2000). The remaining lifetimes beyond this point are an unimportant correction. Stars older than this lifetime are modeled as remnants using the white dwarf initial/final mass relation  $M_{\text{WD}} = 0.109M + 0.394M_\odot$  of Kalirai et al. (2008) for  $M_{\text{NS}} < 8M_\odot$ , a neutron star mass of  $1.35M_\odot$  for  $8M_\odot < M < 40M_\odot$ , and a black hole mass of  $5M_\odot$  for masses  $M_{\text{BH}} > 40M_\odot$ . Using the initial mass function from Chabrier (2003), a combination of a log-normal distribution at low mass and a power-law at high mass covering  $0.01M_\odot < M < 100M_\odot$ , the mean mass is

$$\langle M \rangle \simeq (0.20 + 0.03 \log(t/10 \text{ Gyr})) M_\odot \quad (13)$$

for any reasonable population age  $t$ . Age has little effect because the high mass stars which evolve on these time scales make a limited contribution to the mean mass, and the mass scale beyond which stars have evolved depends weakly on age ( $M_{\text{evolve}} \simeq (t/11 \text{ Gyr})^{0.31} M_\odot$  for the Hurley et al. (2000) models). Changes in the white dwarf mass relations also have little effect. For  $M_{\text{WD}} = aM + b$ ,



the sensitivity is  $\delta\langle M \rangle \sim 0.08\delta a M_\odot + 0.05\delta b$ , where Kalirai et al. (2008) estimate uncertainties of  $\delta a \simeq 0.007$  and  $\delta b \simeq 0.025 M_\odot$ . Similarly, changes in the masses of neutron stars and black holes have negligible effects on the mean mass, with  $\delta\langle M \rangle \simeq 0.0035\Delta M_{\text{NS}}$  and  $\simeq 0.00034\Delta M_{\text{BH}}$ , respectively, and the same holds true for the masses defining the boundaries between remnant types. Even giving all stars a binary companion with the secondary mass ratio uniformly distributed from  $1/50 < M_2/M_1 < 1$  affects the mean mass little, roughly  $0.05 M_\odot$ . The sense of the effect depends on the size distribution of the binaries. Very wide binaries act like independent stars and so lower the mean, while very close binaries act like a single, higher mass star and hence raise the (effective) mean.

Thus, only changes in the actual initial mass function can significantly alter the expected mean mass. The Chabrier (2003) mass function converges to low masses, so extending the mass range downwards to  $0.001 M_\odot$  from  $0.01 M_\odot$  reduces the mean mass by only  $0.02 M_\odot$ . Significant changes require a mass function converging more slowly at lower masses or adding entirely new populations. For example, if we instead use a Kroupa et al. (1993) mass function, the results are very sensitive to the low mass cutoff because the mass function is a rising power law ( $\propto M^{-1.3}$ ) to low masses. For minimum masses of  $0.08 M_\odot$  and  $0.01 M_\odot$  the mean masses are  $0.32 M_\odot$  and  $0.15 M_\odot$  respectively. Figure 8 compares these mass functions to our model simple power law model with  $\langle M \rangle = 0.3 M_\odot$  and  $M_{\text{max}}/M_{\text{min}} = 50$ , to show that our maximum likelihood model is in good agreement with expectations for normal stellar populations.

We can also compare this mean mass to microlensing measurements made in our own Galaxy and in other quasar microlensing studies of other lenses. The Massive Compact Halo Object (MACHO) survey measured the most likely mass range of compact objects in the Milky Way Halo to be  $0.15 < M/M_\odot < 0.9$  depending on the halo model Alcock et al. (2000), although these results are broadly questioned (e.g. see Tisserand et al. 2007; Wyrzykowski et al. 2009). Estimates for the Galactic bulge are probably more relevant for comparison to Q2237. Han & Gould (1996) determined that a power law mass function  $dN/dM \propto M^{-2.1}$  for  $M > 0.04 M_\odot$  was the best fit to a sample of 51 MACHO Galactic bulge microlensing events. This corresponds to a mean mass of  $0.19 M_\odot$  assuming a maximum mass of  $10 M_\odot$ . Grenacher et al. (1999) studied the first 41 MACHO bulge events toward Baade’s windows and found a mean mass of  $0.09 M_\odot$  (0.129) for bulge (disk) lenses. They assumed a Salpeter mass function in the range 1–10  $M_\odot$  and fit for the best slope and minimum mass below  $1 M_\odot$ . Calchi Novati et al. (2008) found a very similar result. Outside our Galaxy, the only limits aside from those for Q2237 are those for the doubly imaged quasar Q 0957+561 by Schmidt & Wambsganss (1998), who found a weak lower limit of  $\langle M \rangle \gtrsim 0.001 M_\odot$ .

#### 4.3. Magnification Offsets

We can also try to measure the relative mean magnification offsets between each of the images. In our models we do not constrain the mean magnification ratios of the images to closely match the predictions

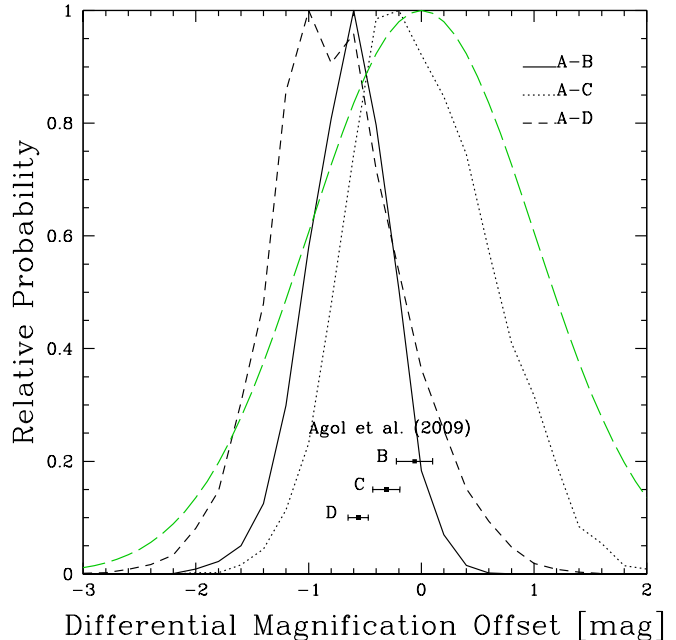


FIG. 9.— Differential magnification offsets,  $\Delta\mu_i$  for A–B, A–C, and A–D. The extinction measurements by Agol et al. (2009) are shown relative to image A. The dashed curve is the prior we used on the magnification offsets.

of the lens model, since differential dust extinction (e.g. Falco et al. 1999; Eigenbrod et al. 2008a; Agol et al. 2009), undetected substructure (e.g. Mao & Schneider 1998; Metcalf & Madau 2001; Kochanek & Dalal 2004; Vegetti et al. 2009), bad “macro” models of the lens magnification, and contamination of the light curves by light from the lens or host galaxy can also change the relative brightnesses of the images. We allow  $\Delta\mu_i$  (Equation 4) to be optimized for each fit, subject to a Gaussian prior with a 1.0 magnitude dispersion. In Figure 9 we show the posterior probability distributions for these differential offsets. For an infinitely long light curve, these offsets will converge to zero in the absence of any systematic problems. The differential offsets between A–B and A–D show weak evidence for offsets, but there is surprisingly little convergence in their values. For comparison, Agol et al. (2009) used the flux ratios of the quasar broad lines as compared to the continuum from Eigenbrod et al. (2008a) to estimate the extinction of the images relative to A. They found  $\Delta E(B - V) = 0.02 \pm 0.05, 0.10 \pm 0.04$ , and  $0.18 \pm 0.03$  for images B, C, and D respectively. Figure 9 shows these estimates assuming a  $R_V = 3.1$  extinction curve. There is some correlation between our estimates and these shifts, but our estimates are simply too uncertain to draw any conclusions. We experimented with forcing our trials to match the extinction estimates of Agol et al. (2009) by multiplying the probability of each trial by a Gaussian model of these extinction estimates. We found no significant influence on any other parameter distribution. Dai et al. (2009) reached a similar conclusion in their analysis for RXJ1131–1231.

## 5. DISCUSSION

By including the random motions of the stars, we can now use microlensing to study the peculiar velocity of the

lens galaxy and to estimate mean stellar masses and potentially the stellar mass functions with fewer systematic uncertainties. In particular, we find a clear preference for the direction of motion of the lens galaxy. In fact, as we use a less restrictive velocity prior, the direction of motion is better constrained since faster speeds are allowed. We cannot however, determine the speed without additional priors. If we assume a mean stellar mass of  $\langle M \rangle = 0.3M_{\odot}$ , we find that the peculiar velocity is  $v_t < 486 \text{ km s}^{-1}$  which is consistent with the other estimates (Wyithe et al. 1999; Gil-Merino et al. 2005) but more fully includes all the physical uncertainties. It should not be surprising that we can determine a dimensionless quantity, the direction of motion, better than the dimensional speed, given the basic problem of microlensing that all observables are in  $\sqrt{\langle M \rangle} \text{ cm}$ . Very roughly (see Figure 1), the preferred direction has images A and B moving more closely to perpendicular to the ridges of the magnification patterns created by the shear and images C and D are moving more parallel to the shear direction. This is consistent with the variability of A/B compared to C/D. We included the parallax effects of the Earth's orbit, and the results weakly favored its inclusion.

We had hoped that modeling the random motions would be more of a help in breaking these degeneracies by setting a physical scale. This is probably true for low mean masses  $\langle M \rangle$ . For fixed variability amplitudes, reproducing the light curves with a low mean mass requires small physical velocities, while high mean masses require high velocities. Adding the stellar motions at their observed dispersion eliminates low mass solutions independent of the unknown peculiar velocities by setting a floor to the velocity scale. High mass solutions need peculiar velocities,  $\sigma$ , that are larger than the stellar motions,  $\sigma_*$ , and so are only constrained by the priors on the peculiar velocities. Essentially, the dynamic patterns act like static patterns once  $\sigma_* < \sigma$ , and we recover the familiar degeneracies of static patterns. Thus, our correct treatment of the stellar motions constrains low mass but not high mass solutions in the absence of a peculiar velocity prior. With a well-defined cosmological prior on  $\sigma$  (Tinker, Wetzel, & Zehavi 2009), we find that  $0.12 \leq \langle M/M_{\odot} \rangle \leq 1.94$  at 68% confidence, demonstrating that the microlensing objects are typical of stellar populations and their remnants. This mass range is consistent with expectations for normal stellar populations (see §4.2), but not tightly constraining.

We largely ignore the macro magnifications predicted by the mass distribution of the lens galaxy in our calculations because of their systematic uncertainties. However some recent studies have made use of this information by analyzing image pairs straddling a critical curve which should have the same magnification (Floyd et al. 2009; Bate et al. 2008). A concern is that the macro magnification may be affected by undetected substructure, differential extinction, or contamination by the lens or host galaxy. In our standard analysis we use the AC signal and largely discard the DC signal by not tightly constraining the mean magnification. Given sufficiently long light curves, the results will converge to the true magnification offsets. Even for Q2237, with its decade long OGLE light curves, the data are not sampling long enough paths across the patterns (see Figure 1) to show

convergence. At present, the distribution of differential mean magnification offsets are too broad (Figure 9) to tightly constrain any systematic magnification offsets. Fortunately, our results for the other physical parameters are little affected by whether we allow these offsets to vary or constrain them with the extinction estimates of Agol et al. (2009).

Finally, we show for the first time that microlensing variability in a lens gives the same results when analyzing different portions of its light curve. The analysis of light curves LC1 and LC2, corresponding to the 1st and 2nd halves of the 11 year OGLE monitoring period, lead to statistically consistent distributions for every parameter we consider. This both confirms our ability to measure parameters and gives us tighter constraints after combining the results. It would be computationally challenging to analyze the full light curve simultaneously because it becomes (exponentially?) harder to fit longer light curves. However, such full analyses are likely needed for some quantities, particularly the magnification offsets, to converge.

In the future we will likely include binaries, even though their effect is not likely to influence the results other than interpreting the meaning of the mean stellar mass (by up to  $0.05M_{\odot}$ , as discussed in §4). However, like the projection of our motion relative to the CMB, the streaming velocities in Q2237 are small compared to the peculiar velocities, and so are do little to break the degeneracy. The effects of streaming velocities will be seen most strongly in true disk lenses (none are known, except, potentially PMN J2004–1349, Winn et al. (2001)), or in lenses such as Q J0158–4325 (see Morgan et al. (2008) for a microlensing analysis of this active system) lying close to the equator of the CMB dipole, which will have the full  $369 \text{ km s}^{-1}$  dipole motion (Hinshaw et al. 2009). These CMB equatorial lenses should also show significantly shorter microlensing variability time scales. Detecting this effect would be an independent confirmation of the kinematic origin of the dipole.

Q2237 was a natural first candidate for a full analysis with moving stars because of the excellent OGLE data, short microlensing timescales, and negligible time delays between the images. However, there is no problem extending our approach to analyzing microlensing data with moving stars to any other microlensing analysis. Even if the time delays are unknown, cases with different trial delays could simply be tried sequentially (Morgan et al. 2008). Moreover our method can easily be extended to multi-wavelength data sets to examine the structure of the accretion disk varies with wavelength (Poindexter et al. 2008). The memory requirements would be too great to fit each band simultaneously as in Poindexter et al. (2008), but we can use a modified version of the method Dai et al. (2009) applied to the joint optical and X-ray analysis of RXJ1131–1231. The models are first run on the band with the most epochs. As good fitting trials are found, the starting points, velocities, and  $\chi^2$  matrix are saved. Next, for each successive band, we recompute the light curves corresponding to the epochs and source sizes of the other wavelengths, and the results of these new fits are used to continue the  $\chi^2$  calculation. Since the overall execution times are only modestly longer than using static stars, there is no reason not to use this more physically correct approach.

SP thanks Jason Sawin for discussions on computational techniques. We thank Tyoma Tuntsov for suggesting the inclusion of the parallax effect. We thank Andrew Gould for providing the parallax code. This work

was supported in part by an allocation of computing time from the Ohio Supercomputer Center. This research was

## REFERENCES

- Agol, E., Jones, B., & Blaes, O. 2000, *ApJ*, 545, 657
- Agol, E., Gogarten, S. M., Gorjian, V., & Kimball, A. 2009, *ApJ*, 697, 1010
- Alcock, C., et al. 2000, *ApJ*, 542, 281
- Anguita, T., Schmidt, R. W., Turner, E. L., Wambsganss, J., Webster, R. L., Loomis, K. A., Long, D., & McMillan, R. 2008, *A&A*, 480, 327
- Bate, N. F., Floyd, D. J. E., Webster, R. L., & Wyithe, J. S. B. 2008, *MNRAS*, 391, 1955
- Calchi Novati, S., de Luca, F., Jetzer, P., Mancini, L., & Scarpetta, G. 2008, *A&A*, 480, 723
- Chabrier, G. 2003, *PASP*, 115, 763
- Chartas, G., Kochanek, C. S., Dai, X., Poindexter, S., & Garmire, G. 2009, *ApJ*, 693, 174
- Congdon, A. B., Keeton, C. R., & Osmer, S. J. 2007, *MNRAS*, 376, 263
- Corrigan, R. T., et al. 1991, *AJ*, 102, 34
- Dai, X., Kochanek, C. S., Chartas, G., Kozlowski, S., Morgan, C. W., Garmire, G., & Agol, E. 2009, *arXiv:0906.4342*
- Eigenbrod, A., Courbin, F., Sluse, D., Meylan, G., & Agol, E. 2008a, *A&A*, 480, 647
- Eigenbrod, A., Courbin, F., Meylan, G., Agol, E., Anguita, T., Schmidt, R. W., & Wambsganss, J. 2008b, *A&A*, 490, 933
- Falco, E. E., et al. 1999, *ApJ*, 523, 617
- Fischer, D. A., & Marcy, G. W. 1992, *ApJ*, 396, 178
- Floyd, D. J. E., Bate, N. F., & Webster, R. L. 2009, *arXiv:0905.2651*
- Foltz, C. B., Hewett, P. C., Webster, R. L., & Lewis, G. F. 1992, *ApJ*, 386, L43
- Gil-Merino, R., Wambsganss, J., Goicoechea, L. J., & Lewis, G. F. 2005, *A&A*, 432, 83
- Gil-Merino, R., & Lewis, G. F. 2005, *A&A*, 437, L15
- Gould, A. 1995, *ApJ*, 444, 556
- Gould, A. 2000, *ApJ*, 535, 928
- Grenacher, L., Jetzer, P., Str ssle, M., & de Paolis, F. 1999, *A&A*, 351, 775
- Han, C., & Gould, A. 1996, *ApJ*, 467, 540
- Hinshaw, G., et al. 2009, *ApJS*, 180, 225
- Huchra, J., Gorenstein, M., Kent, S., Shapiro, I., Smith, G., Horine, E., & Perley, R. 1985, *AJ*, 90, 691
- Hurley, J. R., Pols, O. R., & Tout, C. A. 2000, *MNRAS*, 315, 543
- Irwin, M. J., Webster, R. L., Hewett, P. C., Corrigan, R. T., & Jedrzejewski, R. I. 1989, *AJ*, 98, 1989
- Kalirai, J. S., Hansen, B. M. S., Kelson, D. D., Reitzel, D. B., Rich, R. M., & Richer, H. B. 2008, *ApJ*, 676, 594
- Kayser, R., Refsdal, S., & Stabell, R. 1986, *A&A*, 166, 36
- Kayser, R., & Refsdal, S. 1989, *Nature*, 338, 745
- Kochanek, C. S., 2004, *ApJ*, 605, 58
- Kochanek, C. S., & Dalal, N. 2004, *ApJ*, 610, 69
- Kochanek, C. S., Dai, X., Morgan, C., Morgan, N., Poindexter, S., & Chartas, G., 2007, in *Statistical Challenges in Modern Astronomy IV in Statistical Challenges in Modern Astronomy IV* G. J. Babu and E. D. Feigelson, eds., (Astron. Soc. Pacific: San Francisco), [astro-ph/0609112]
- Keeton, C. R., 2001, *astro-ph/0102340*
- Kent, S. M., & Falco, E. E. 1988, *AJ*, 96, 1570
- Kroupa, P., Tout, C. A., & Gilmore, G. 1993, *MNRAS*, 262, 545
- Kundic, T., & Wambsganss, J. 1993, *ApJ*, 404, 455
- Lee, D.-W., Surdej, J., Moreau, O., Libbrecht, C., & Claeskens, J. -. 2005, *arXiv:astro-ph/0503018*
- Lewis, G. F., & Irwin, M. J. 1996, *MNRAS*, 283, 225
- Metcalf, R. B., & Madau, P. 2001, *ApJ*, 563, 9
- Mao, S., & Schneider, P. 1998, *MNRAS*, 295, 587
- Mediavilla, E., et al. 2009, *ApJ*, 706, 1451
- Mihov, B. M. 2001, *A&A*, 370, 43
- Morgan, C. W., Kochanek, C. S., Dai, X., Morgan, N. D., & Falco, E. E. 2008, *ApJ*, 689, 755
- Morgan, C. W., Eyler, M. E., Kochanek, C. S., Morgan, N. D., Falco, E. E., Vuissoz, C., Courbin, F., & Meylan, G. 2008, *ApJ*, 676, 80
- Morgan, C., Kochanek, C. S., Morgan, N. D., Falco, E. E., 2007, *ArXiv e-prints*, 707, *arXiv:0707.0305*, in prep.
- Mortonson, M. J., Schechter, P. L., Wambsganss, J., 2005, *ApJ*, 628, 594
- Mosquera, A. M., Mu  oz, J. A., & Mediavilla, E. 2009, *ApJ*, 691, 1292
- Paczynski, B. 1986, *ApJ*, 301, 503
- Poindexter, S., Morgan, N., & Kochanek, C. S. 2008, *ApJ*, 673, 34
- Poindexter, S., & Kochanek, C. S. 2009, in prep.
- Pooley, D., Blackburne, J. A., Rappaport, S., & Schechter, P. L. 2007, *ApJ*, 661, 19
- Pooley, D., Rappaport, S., Blackburne, J., Schechter, P. L., Schwab, J., & Wambsganss, J. 2009, *ApJ*, 697, 1892
- Rix, H.-W., Schneider, D. P., & Bahcall, J. N. 1992, *AJ*, 104, 959
- Schmidt, R., & Wambsganss, J. 1998, *A&A*, 335, 379
- Shakura, N. I. & Sunyaev, R. A., 1973, *A&A*, 24, 337
- Schechter, P. L., Wambsganss, J., & Lewis, G. F. 2004, *ApJ*, 613, 77
- Schneider, D. P., Turner, E. L., Gunn, J. E., Hewitt, J. N., Schmidt, M., & Lawrence, C. R. 1988, *AJ*, 96, 1755
- Schramm, T., Kayser, R., Chang, K., Nieser, L., & Refsdal, S. 1993, *A&A*, 268, 350
- Tinker, J. L. & Wetzel, A. R., Zehavi, I., 2009, in prep.
- Tisserand, P., et al. 2007, *A&A*, 469, 387
- Trott, C. M., & Webster, R. L. 2002, *MNRAS*, 334, 621
- Trott, C. M., Treu, T., Koopmans, L. V. E., & Webster, R. L. 2008, *arXiv:0812.0748*
- Tuntsov, A. V., Walker, M. A., & Lewis, G. F. 2004, *MNRAS*, 352, 125
- Udalski, A., et al. 2006, *Acta Astronomica*, 56, 293
- van de Ven, G., Falcon-Barroso, J., McDermid, R. M., Cappellari, M., Miller, B. W., & de Zeeuw, P. T. 2008, *arXiv:0807.4175*
- Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T., & Gavazzi, R. 2009, *arXiv:0910.0760*
- Wambsganss, J., Paczynski, B., & Katz, N. 1990, *ApJ*, 352, 407
- Wambsganss, J., & Paczynski, B. 1994, *AJ*, 108, 1156
- Wambsganss, J., & Kundic, T. 1995, *ApJ*, 450, 19
- Wambsganss, J. 2006, in *Gravitational Lensing: Strong Weak and Micro, Saas-Fee Advanced Course 33*, G. Meylan, P. North, P. Jetzer, eds., (Springer: Berlin) 453, [astro-ph/0604278]
- Winn, J. N., Hewitt, J. N., Patnaik, A. R., Schechter, P. L., Schommer, R. A., L  pez, S., Maza, J., & Wachter, S. 2001, *AJ*, 121, 1223
- Wyithe, J. S. B., Webster, R. L., & Turner, E. L. 1999, *MNRAS*, 309, 261
- Wyithe, J. S. B., Webster, R. L., & Turner, E. L. 2000, *MNRAS*, 312, 843
- Wyithe, J. S. B., Webster, R. L., & Turner, E. L. 2000, *MNRAS*, 315, 51
- Wyithe, J. S. B., & Turner, E. L. 2001, *MNRAS*, 320, 21
- Wyrzykowski, L., et al. 2009, *MNRAS*, 397, 1228